

# ACENTO E PRODUÇÃO DE FALA

Este capítulo apresenta abordagens teóricas e experimentais do acento sob um ponto de vista da produção da fala. Primeiramente, iremos delimitar o fenômeno do acento lexical, apresentando observações de ordem teórica e terminológica. Traçaremos então um panorama histórico sobre como evoluiu a concepção do acento em algumas teorias linguísticas, do estruturalismo aos modelos gestuais, passando pelo gerativismo e modelos fonológicos não lineares. Em seguida, serão descritas as propriedades acústicas e fonatórias geralmente associadas ao fenômeno de proeminência acentual. Na subseção seguinte, resumiremos os aspectos consensuais e os principais debates enfrentados pelas teorias fonológicas sobre o acento lexical no português brasileiro. Por fim, apresentaremos alguns estudos acústicos do acento no português brasileiro.

### 3.1 DELIMITAÇÃO DO FENÔMENO ACENTUAL

Tradicionalmente, os estudos sobre a dimensão sonora da fala humana são divididos em Fonética e Fonologia. Enquanto a primeira se dedica ao detalhamento dos aspectos físicos dos sons da fala – em termos de suas propriedades acústicas, aerodinâmicas, fonatórias, articulatórias, auditivas –, a segunda se volta para o funcionamento e a organização de tais sons – em termos de estruturas e mecanismos

cognitivos subjacentes. Trata-se, portanto, de uma oposição quanto à perspectiva de análise adotada<sup>1</sup>. A área da Prosódia, que comporta tanto uma abordagem do tipo fonético quanto fonológico, se define antes pela especificidade de seu tema: dedica-se ao estudo de fenômenos não segmentais, geralmente atribuídos à “voz”, como ritmo, entoação e acento. Ritmo (*rhythm*) se refere aos padrões que surgem na fala “[...] out of the periodic recurrence of some sort of movement, producing an expectation that the regularity of succession will continue.” (ABERCROMBIE, 1968, p. 96). A entoação (*intonation*) é um dos aspectos que contribuem com a organização do ritmo da língua e diz respeito ao uso da variação da altura para transmitir “[...] sentence-level pragmatic meanings in a linguistically structured way.” (LADD, 2008, p. 4). A definição de acento (*stress*), fenômeno que também contribui para as características rítmicas de uma língua e tema central deste trabalho, se abre a variadas perspectivas. De um lado, há definições que enfocam a interação do acento com o plano entoacional: “Stress is an abstract positional property of linguistic units which denotes the position at which a non-contrastive accent [i.e. intonational prominence] can occur. The stress position is determined by the structure of words and phrases.” (HIRST; DI CRISTO, 1998, p. 86). Por outro lado, em uma perspectiva fonética, como a de Hogg e McCully (1987, p. 1-2), “[...] stress is associated with an increase in respiratory activity on the part of the speaker which causes an increase in any one or other of the following: vowel length, pitch, loudness.” Note-se que a definição de Hog e McCully relaciona um mecanismo aerodinâmico com grandezas perceptuais (*length*, *pitch* e *loudness*, que preferimos não traduzir, para fugir a problemas terminológicos), definidas como correlatos das grandezas acústicas duração, frequência fundamental (F0) e intensidade acústica. No capítulo 4, que trata da percepção de padrões acústicos, retomaremos a relação entre os três pares de propriedades mencionados acima.

Para alguns autores, ainda, acento pode ser definido simplesmente como proeminência silábica:

In the past, the word ‘stress’ has been used in different and confusing ways. It has sometimes been used simply to refer to syllables (or vowels) made prominent for linguistic purposes, either in words or in sentences. But stress has also often been used to mean ‘breath-force or loudness’ the implication being that this is the principal means whereby syllables are made prominent. This second type of usage is misguided since [...] loudness generally plays a minor role in producing prominences. [...]

<sup>1</sup> A divisão da dimensão sonora em Fonética e Fonologia é controversa tanto em sua separação meramente metodológica, em vista da interferência entre os dois níveis de análise (não raro o fonólogo toma decisões com base em detalhes físicos dos sons e o foneticista, com base em seu funcionamento), quanto em sua separação orgânica respectivamente no módulo de implementação e no módulo de planejamento da linguagem, sendo questionada a oposição entre a capacidade da linguagem e seu uso (cf. cap. 2).

I shall use the term STRESS to mean ‘prominence’, however such prominence is achieved. (CRUTTENDEN, 1997, p. 13).

A definição de “acento” ou “acento lexical” que será usada neste estudo buscará integrar as três perspectivas apresentadas acima. Contudo, Hyman (1977) já apontava para a necessidade de se conhecer mais a respeito do acento. Anos se passaram e resta ainda controvérsia a respeito da própria definição conceitual e terminológica do fenômeno, sem que se tenha chegado ainda a uma posição consensual sobre sua natureza, como se vê em texto relativamente recente do mesmo autor (HYMAN, 2006). Por exemplo, ainda que a oposição entre acento e tom seja, de um modo geral, aceita e definida com alguma facilidade, surgem problemas ao se diferenciar acento no que Hyman chama de *stress-accent* e *pitch-accent*, devido à semelhança da função que desempenham na fonologia das línguas.

Nota-se que a Prosódia enfrenta ainda muitos problemas quanto a definições terminológicas, como retomaremos adiante, em especial, problemas referentes à tradução dos conceitos prosódicos, à multiplicidade de nomenclaturas e ainda à classificação dos elementos prosódicos. Para fins de clareza, iremos apresentar o termo original em inglês sempre que um novo conceito for introduzido no texto. Destacamos ainda que será evitado o termo “tonicidade” como referência ao fenômeno do acento, uma vez que etimologicamente “tônico” pode passar uma falsa ideia de que se trata de uma marcação acentual baseada em F0. Como veremos mais adiante, o acento apresenta simultaneamente mais de um correlato acústico possível. Exceção será feita para denominar as sílabas átonas em dois tipos: as que antecedem a sílaba acentuada (“pretônicas”) ou que a sucedem (“postônicas”), pela falta de termos substitutivos eficientes. Além disso, o termo “acento” que será utilizado ao longo deste estudo corresponde à tradução de *stress*, e não de *accent*, sendo o último geralmente empregado com referência ao destaque conferido a palavra ou sílaba em níveis prosódicos superiores, como o sintagma, a frase ou o enunciado (CRYSTAL, 1995).

Em algumas línguas, o acento lexical pode recair sobre mais de uma sílaba na mesma palavra, havendo uma gradação no nível de proeminência em cada uma das sílabas acentuadas. Nesse caso, diz-se que esta língua apresenta acento primário (que incide sobre a sílaba mais proeminente) e também secundário (incidente sobre as sílabas com menor grau de proeminência).

[...] in any word [...] one and only one vowel receives the strongest stress of the word. We should call this stress *primary stress*. [...] the vowels in any constituent which are stressed but do not receive the strongest stress of the constituent are said to bear *secondary stress*. [...] Only vowels which are initially assigned primary stress can, in some larger context, bear secondary stress (HOGG; MCCULLY, 1987, p. 6).

Iremos restringir nossa análise somente ao acento lexical primário do PB, visto que o acento secundário é um fenômeno complexo, que mereceria um estudo à parte. Abordagens aprofundadas do assunto podem ser encontradas e.g. em Collischonn (1994) e Arantes (2010a). No presente trabalho, tanto “acento lexical” quanto “acento” serão empregados como equivalentes a “acento lexical primário”.

Na perspectiva que Abercrombie (1968) denomina de Fonética Clássica, os componentes acústicos da fala foram divididos em segmentais e não segmentais (ou prosódicos), sendo os últimos subdivididos em qualidade vocal e dinâmica vocal.

The strand consisting of the segmental features of an utterance is made up of complex auditory qualities which are in fairly rapid fluctuation, reflecting the rapid succession of movements of the articulators. The strand of voice dynamics also consists of features which fluctuate in auditory quality, but considerably more slowly. Many of them are linked to the syllable- and stress-producing processes [...], and they are closely related to [...] aspects of sound [...] such as pitch loudness, tempo ad rhythm. The strand consisting of features of voice quality in contrast to the two preceding, has a quasi-permanent character: it remains constant over relatively long stretches of time, and fluctuation here is much less apparent. (ABERCROMBIE, 1968, p. 90).

Os fenômenos prosódicos são, via de regra, realizados por meio de três parâmetros acústicos básicos – F0, intensidade e duração –, podendo haver sobreposição entre os três parâmetros na implementação de diferentes fenômenos (LADEFOGED, 1996; LAVER, 1994). Em uma mesma língua, por exemplo, acento e entoação, ainda que implementados independentemente, podem se valer de mecanismos acústico-fonatórios similares. Além disso, a interação entre os mecanismos acústicos de implementação do acento apresenta especificidades em cada língua. Por essa razão, uma definição puramente fonética do acento, como “[t]he use of extra respiratory energy during a syllable” (LADEFOGED, 1982, p. 285), não caracteriza todo o fenômeno. Mostra-se necessário descrevê-lo considerando também suas características funcionais. O presente estudo busca oferecer contribuições para a descrição do acento sob um ponto de vista funcional e experimental.

## 3.2 O ESTUDO DO ACENTO AO LONGO DO DESENVOLVIMENTO HISTÓRICO DAS TEORIAS FONOLÓGICAS

### 3.2.1 Estruturalismo

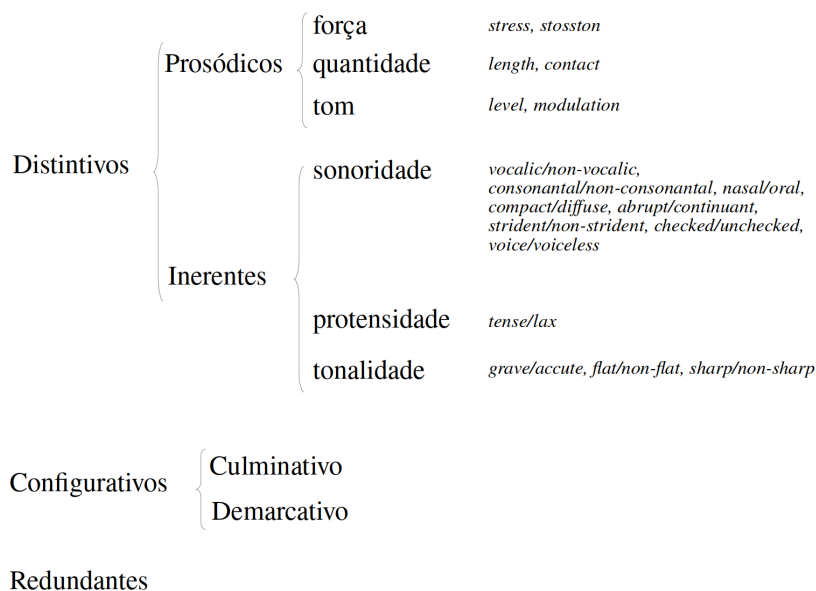
A tradição estruturalista inaugurou a descrição dos sons das línguas do mundo por meio da combinação de propriedades básicas, os traços. Tais propriedades foram primeiramente elaboradas nos estudos fonéticos, como forma de descrever os segmentos

sonoros físicos (fones) de acordo com suas características articulatórias. Foram então reelaboradas em uma perspectiva fonológica e passaram a contemplar também características acústicas, com fins de definir os segmentos sonoros abstratos (fonemas). Construiu-se um sistema, que se supunha universal, de traços distintivos, definidos como as unidades mínimas no plano sonoro<sup>2</sup>. A partir dos traços, seriam obtidas, combinatoriamente, todas as estruturas sonoras da língua: fonemas, sílabas, palavras.

The distinctive features are aligned into simultaneous bundles called phonemes; phonemes are concatenated into sequences; the elementary pattern underlying any grouping of phonemes is the SYLLABLE. The phonemic structure of the syllable is determined by a set of rules, and any sequence is based on the regular recurrence of this constructive model. A FREE FORM (a sequence, separable by means of pauses) must contain an integral number of syllables. [...] The pivotal principle of syllable structure is the contrast of successive features within the syllable. One part of the syllable stands out against the others. It is mainly the contrast vowel vs. consonant which is used to render one part of the syllable more prominent. (JAKOBSON; HALLE, 1971, p. 31).

A seguir a FIG. 3 esquematiza, a título de ilustração, a tipologia de traços sonoros proposta por Jakobson e Halle (1971)<sup>3</sup>. Note-se que há traços que se referem a propriedades acústicas, articulatórias e funcionais dos segmentos.

**FIGURA 3** – Sistema de traços fonológicos proposto em Jakobson e Halle (1971)



<sup>2</sup> Para diferentes propostas de sistemas de traços fonológicos, cf. Kenstowicz (1993).

<sup>3</sup> Os nomes dos traços não foram traduzidos, para evitar problemas terminológicos não pertinentes a este estudo. Os pares de traços agrupados em cada categoria – força, quantidade e tom – correspondem, de cima para baixo, ao nível intersilábico e intrasilábico.

Traços distintivos do tipo prosódico seriam propriedades exclusivas de segmentos que exercem função de núcleo da sílaba (geralmente vogais), enquanto os traços do tipo inerente abrangeriam todos os tipos de segmentos, independente de seu papel na estrutura silábica. Os autores agrupam os traços prosódicos nas categorias de força, quantidade e tom, correspondentes, no plano perceptual, respectivamente a *loudness*, duração relativa e *pitch*, e, no plano acústico, a intensidade (amplitude), duração absoluta (tempo) e frequência.

Os autores argumentam que, ao contrário dos traços inerentes, tidos como absolutos, os traços prosódicos implicam no estabelecimento de contraste entre variáveis em determinado domínio: por exemplo, uma sílaba é considerada mais proeminente em comparação a outra sílaba da mesma palavra.

A prosodic feature involves two coordinates: on the one hand, polar terms such as high or low register, rising and falling pitch, or long and short, all may appear, *ceteris paribus*, in the same position within the sequence; thus the speaker selectively uses and the listener selectively apprehends one of the two alternatives and identifies the chosen alternative in relation to the rejected one. [...] On the other hand, both polar terms are fully recognizable only when both of them are present in the given sequence, so that the speaker effects and the listener perceives their contrast. Thus both alternative of a prosodic feature co-exist in the code as two terms of an opposition and, moreover, co-occur and produce a contrast within the message. (JAKOBSON; HALLE, 1971, p. 37).

Traços prosódicos são, portanto, propriedades sempre relativas e, consequentemente, variáveis em termos de sua magnitude absoluta, seja entre falantes diferentes ou entre locuções de um mesmo falante.

Outro aspecto destacado por Jakobson e Halle diz respeito à função do acento, especificada por meio dos traços configurativos que poderiam interagir com o traço distintivo *stress*. A função culminativa diz respeito à presença de apenas um pico máximo de proeminência local (seja em uma palavra, geralmente de conteúdo, ou em um grupo formado por palavras funcionais anexadas a palavras de conteúdo<sup>4</sup>), que serve como âncora para a estrutura prosódica (PIERREHUMBERT, 1980; PIERREHUMBERT; BECKMAN, 1988). Já a função demarcativa do acento se refere à sinalização de fronteiras gramaticais, seja de morfemas, palavras, ou grupos de palavras.<sup>5</sup>

<sup>4</sup> Compare, e.g., a presença de duas proeminências acentuais nos sintagmas “**bela casa**” e apenas uma em “a **casa**”. (Ao longo deste estudo, quando relevante, sílabas acentuadas serão marcadas por negrito e representação fonética dos sons será dada entre colchetes: “[ ]”).

<sup>5</sup> No português, o acento lexical acontece sempre em uma das três últimas sílabas da palavra (e.g. **cará**, **cara**, **máscara**) ou na quarta última, em casos específicos de epêntese e próclise (e.g. **rí**[i]mico e **chamáv**amo-nos).

Wherever there is a contrast of stressed and unstressed syllables, stress is always used as a configurative, namely culminative feature, whereas length never assumes this function. The culminative function of stress is regularly combined either with the other variety of configurative functions, demarcation, or with the distinctive function. Languages in which both length and stress appear as mutually independent distinctive features are quite exceptional, and if the stress is distinctive, it is frequently supplemented by redundant length. The observation of force and quantity features in their intersyllabic variety seem to indicate that the prosodic distinctive features utilizing intensity and those utilizing time tend to merge. (JAKOBSON; HALLE, 1971, p. 36).

Comopode-se observar, os autores ressaltam que, em várias línguas, como consequência da coocorrência redundante dos parâmetros acústicos de intensidade e duração na realização do acento, pode haver fusão entre essas propriedades. Como retomaremos na seção seguinte, o acento lexical como o do português mostra-se um caso complexo para a caracterização das unidades sonoras por meio de traços: é ao mesmo tempo um traço do tipo distintivo prosódico (*stress* e *length*), do tipo configurativo (culminativo e demarcativo) e do tipo redundante (dada a fusão entre os traços prosódicos). Além disso, sua função configurativa como âncora prosódica e delimitador de palavras é um indicativo de que sua abrangência extrapola o domínio de segmentos individuais, como veremos adiante.

### 3.2.2 Gerativismo e modelos não lineares

O gerativismo não apresentou mudanças significativas com relação ao estruturalismo quanto à concepção de acento. Chomsky e Halle (1968), considerando a fonologia do inglês, definem o acento como um traço [+– stress] atribuído à vogal por meio de uma regra específica da língua (*Nuclear Stress Rule*) e que depende da estrutura sintática. No entanto, a abordagem baseada em traços, com atribuição do acento por meio de algoritmos específicos, como a defendida pela fonologia gerativista, logo se deparou com problemas em lidar com características prosódicas da fala.

Alguns dos modelos fonológicos que surgiram nas décadas de 70 e 80, ditos Modelos Não Lineares<sup>6</sup>, buscaram solucionar problemas de interação dos segmentos com a prosódia. Uma das grandes inovações dessa época foi a proposta de interpretar o plano sonoro em uma estrutura multi-linear. A Fonologia Autossegmental desenvolveu a ideia de que os elementos prosódicos, como tons lexicais e acento, seriam suprasegmentos (GOLDSMITH, 1976) – i.e. ocorreriam em um plano hierárquico superior ao plano segmental, uma vez que podem operar sobre uma cadeia de segmentos. O modelo autossegmental ganhou elaboração teórica com Geometria de Traços

<sup>6</sup> Os modelos fonológicos chamados de não lineares se comprometem com os axiomas fundamentais da Gramática Gerativa, buscando contribuir com a adequabilidade desse programa de pesquisa na área do componente sonoro.



(CLEMENTS, 1985), em que a matriz de traços definidores de um segmento passa a ser vista não simplesmente como uma lista desordenada de atributos independentes. Pelo contrário, a Geometria de Traços sugere que os traços podem ser agrupados hierarquicamente em nós referentes a pontos articulatórios. Os traços seriam atados a um nó de raiz, sendo este dominado por uma camada temporal, representando a duração do segmento. Segmentos são associados aos constituintes silábicos – *onset*, núcleo, coda –, que, por sua vez, se vinculam à sílaba e ao plano de atribuição do acento. Uma das contribuições mais significativas da Fonologia Autossegmental para o estudo do acento, ao lado da noção de não linearidade – i.e., multidimensionalidade – dos suprassegmentos, está relacionada à noção da sílaba como unidade de *status* teórico.

Outro modelo não linear que paralelamente ofereceu avanços para o estudo do acento foi a Fonologia Métrica, que, voltando-se para a estrutura rítmica de palavras e sentenças, conferiu um tratamento unificado ao acento lexical e frasal, como fenômenos da mesma natureza, diferenciando-se apenas por se aplicar a domínios diferentes: “[...] stress patterns reflect, to some extent at least, relations of prominence between constituents. [...] Thus stress assignment in metrical phonology is a matter of denoting the relative prominence of sister constituent.” (HOGG; MCCULLY, 1987, p. 64-65). Liberman e Prince (1977) defendem que o acento não é uma propriedade de segmentos individuais, ou sílabas, como proposto pelo estruturalismo e gerativismo, mas antes reflete uma estruturação rítmica hierárquica responsável por organizar as sílabas, palavras e sentenças de um enunciado. Os autores apontam inconsistências na abordagem do acento como um traço distintivo, argumentando que o traço [*stress*] não se comporta como os outros traços do sistema: os valores do traço [*stress*] são n-ários; a possibilidade de níveis diferentes de acento dentro de um mesmo domínio leva à necessidade de regras de acento secundário, terciário etc., no nível sintagmático, enquanto os outros traços fonéticos são definidos paradigmaticamente; regras acentuais não são locais, como as regras segmentais. Tais problemas seriam solucionados pela abordagem métrica, com o estabelecimento de uma organização binária do acento na estrutura arbórea de constituintes métricos, que definiria relações sintagmáticas, de nível local.

Com a repercussão da abordagem não linear, ênfase foi dada à descrição dos fenômenos prosódicos em termos de seu domínio de ocorrência: há padrões que ocorrem no nível da palavra (*word-level prosody*, como o acento e o tom) e padrões que ocorrem no nível da frase (*phrase-level prosody*, como a entoação), e a organização de tais domínios obedece a princípios hierárquicos estritos, como enfocado na Fonologia Prosódica de Selkirk (1980) e Nespor e Vogel (1986). Ênfase



foi dada à descrição e demonstração da relevância de constituintes prosódicos (HALLE; VERGNAUD, 1987; MCCARTHY; PRINCE, 1999).

Muito se debateu nas teorias formalistas não lineares sobre aspectos paramétricos do acento, específicos de cada língua, especialmente quanto a: número de constituintes por pé; pé com núcleo à esquerda ou à direita; iteratividade e direcionalidade na divisão dos pés; sensibilidade ao peso silábico (cf. HAYES, 1995), com objetivo de desenvolver uma estrutura de princípios prosódicos universais. Alguns desses debates ganharam nova roupagem com o desenvolvimento da Teoria da Otimalidade (Prince; Smolensky, 2002), que apresenta uma nova perspectiva de gramática, rompendo com a noção de regras e representação subjacente. As regras dão lugar ao ranqueamento de restrições violáveis. Parte-se do *output* (que corresponderia à forma de superfície na gerativa) e, por meio da interação de restrições, escolhe-se o candidato ótimo, que será o *input* (o que era a forma subjacente na gerativa).

### 3.2.3 Modelos gestuais

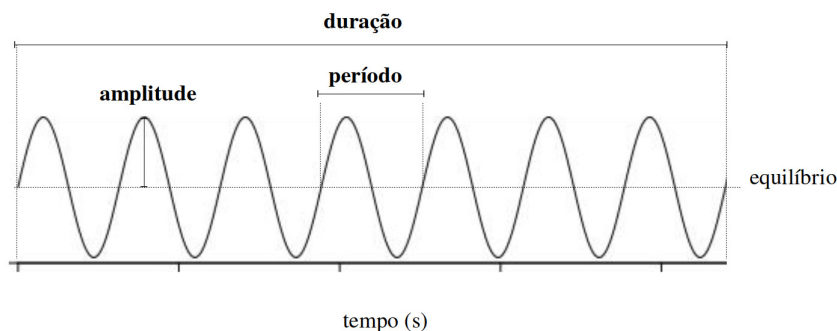
A descrição e modelagem da articulação dos sons da fala ganhou importância crescente nos estudos linguísticos, ao longo das três últimas décadas. A fala passou a ser vista como produto da multimodalidade, dando origem a diversos modelos gestuais em fonologia (e.g. Fonologia Articulatória (BROWMAN; GOLDSTEIN, 1986), *Task Dynamics* (SALTZMAN; MUNHALL, 1989)). Para tais modelos, a fala é construída a partir dos gestos, padrões articulatórios gerados pela movimentação coordenada dos articuladores da fala. Gestos são unidades representacionais definidas dinamicamente, nas dimensões espacial e temporal (GAFOS, 2002). Os modelos gestuais foram primeiramente desenvolvidos para lidar com fenômenos segmentais, mas, na última década, passaram a dispor também de abordagens da estrutura prosódica (BYRD; CHOI, 2010; SALTZMAN et al. 2008). Em comparação com as outras teorias fonológicas, o PB conta com poucos estudos baseados em modelos gestuais. Dos que tratam do acento lexical, podemos mencionar Albano (2001), que distingue a potencialidade de proeminência lexical do acento propriamente dito. A autora atribui a potencialidade de proeminência à magnitude do gesto vocálico, assumindo que a representação das palavras conteria informações sobre as vogais que são acentuáveis: “a vogal final ou pré-final da palavra é lexicalmente reduzida exceto se a fronteira vocabular coincide com certas fronteiras morfológicas” (ALBANO, 2001, p. 150-151). A acentuação se realizaria, então, pela atribuição de proeminência fonética às vogais acentuáveis, o que coloca a abordagem de Albano em posição intermediária entre a hipótese de armazenagem lexical e de implementação por algoritmos.

### 3.3 ASPECTOS FÍSICOS DA PROEMINÊNCIA ACENTUAL

Esta subseção explora a dimensão física do acento, em termos de suas características acústicas e fonatórias. Para tal, é conveniente antes apresentar algumas definições básicas da análise acústica da fala, baseadas em Handel (1993), Stevens (2000), Kent e Read (1992), aos quais remetemos o leitor para uma visão aprofundada sobre o assunto.

O som surge de vibrações formadas por diferenças de pressão que se propagam sob a forma de uma onda. Na Física, as medidas básicas empregadas para descrever uma onda, relevantes para este estudo, são duração, período, amplitude, como representado na FIG. 3. A duração corresponde ao tempo transcorrido do início ao fim de um evento. Em eventos envolvendo um padrão que se repete ao longo do tempo, o período é o tempo transcorrido até que o padrão volte seu valor inicial. A amplitude<sup>7</sup> corresponde ao valor máximo que o padrão atinge. A partir do período e da amplitude, podemos obter as duas outras medidas que, ao lado da duração, são tradicionalmente empregadas na análise da fala: frequência<sup>8</sup> e intensidade. A frequência é o inverso do período e corresponde ao número de repetições de um padrão em determinado espaço de tempo. A relação entre amplitude e intensidade não é tão simples e a intensidade não é tão facilmente visualizável como as demais medidas, mas, para os presentes propósitos, basta dizer que a intensidade é proporcional ao quadrado da amplitude.

FIGURA 4 – Medidas básicas realizadas em uma onda



Note-se que a FIG. 4 corresponde à representação de um som formado por uma única frequência. Os sons da fala, por outro lado, são complexos, pois podem

<sup>7</sup> Amplitude de pico.

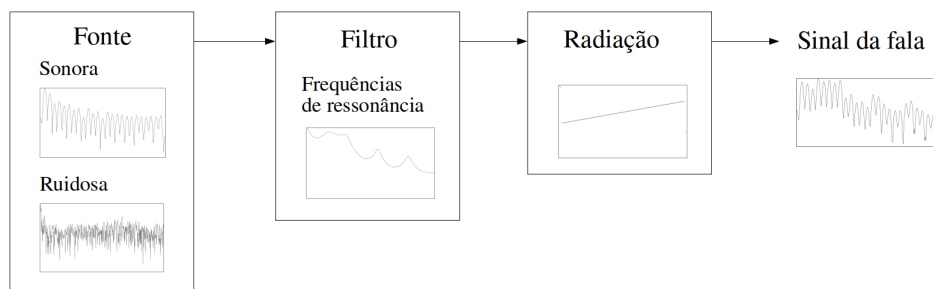
<sup>8</sup> Frequência temporal, que será referida apenas como frequência. Em outra parte deste trabalho, iremos introduzir uma medida de contagem lexical, também chamada de frequência, mas que, para evitar confusões, será especificada como frequência lexical (de ocorrência ou de tipo).

ser decompostos em várias frequências. Os sons que envolvem vozeamento, pela vibração das pregas vocais, apresentam frequências em um padrão específico. A frequência de valor mais baixo é chamada de frequência fundamental e as frequências superiores são chamadas de harmônicos e seus valores são múltiplos inteiros da fundamental. Nos sons da fala, a frequência fundamental, também chamada de  $F_0$ , é determinada pela frequência de vibração das cordas vocais.

A produção de um som requer um posicionamento característico dos articuladores no trato, resultando em uma cavidade frequências de ressonância particulares. Tais frequências de ressonância realizam uma amplificação seletiva de certas regiões de frequências do sinal inicial, produzido pela vibração das cordas vocais. Como consequência, são geradas, no sinal da fala emitido, regiões espectrais com maior concentração de energia, os formantes. Em sons vocálicos, os formantes são infinitos em número, mas apenas os três primeiros ( $F_1$ ,  $F_2$ ,  $F_3$ ) são considerados na análise fonética, pois são suficientes para a caracterização dos sons.

Um modelo clássico em acústica para descrição da fala, conhecido como modelo de fonte-filtro, é mostrado na FIG. 5. Baseia-se no princípio de que os sons da fala resultam de um sinal acústico que é originado de uma fonte, modificado ao passar por uma cavidade de ressonância e, por fim, irradiado.

**FIGURA 5** – Modelo Fonte-Filtro



A fonte é sonora em sons vozeados e corresponde à fonação. A fonte é ruidosa em sons que envolvem alguma perturbação acústica no trato vocal, como aspiração, fricção, explosão. A cavidade de ressonância, que atua como filtro acústico, corresponde ao trato vocal, cuja configuração depende da posição dos articuladores. A etapa de radiação corresponde à emissão da onda sonora ao saírem pelos lábios. O modelo foi amplamente explorado por Fant na Teoria Acústica da Produção da Fala (FANT, 1970), que, ao lado de outros grandes foneticistas de sua época, foi responsável por consolidar métodos experimentais de análise acústica da fala e por descrever os mecanismos acústicos e articulatórios envolvidos na produção da fala. O modelo fonte-filtro descreve de forma

simplificada os processos envolvidos na produção dos sons da fala, procurando estabelecer uma interpretação causal entre eventos acústicos e articulatórios, de forma a ser possível prever certos resultados sonoros a partir de configurações gestuais e vice-versa.

Na descrição e modelagem física dos fenômenos prosódicos, são geralmente considerados, além dos aspectos acústicos, aspectos aerodinâmicos e fonatórios. Procura-se determinar os mecanismos geradores das propriedades acústicas do acento e da entoação, em especial, os correlatos fisiológicos da intensidade e da frequência.

Fant (1981) aponta duas estratégias básicas envolvidas no aumento da intensidade: (1) aumento na pressão subglótica; (2) adução das cordas vocais devido à compressão medial. F0 e intensidade se correlacionam positivamente com a pressão subglótica, cujo aumento pode resultar da ação do sistema respiratório ou do aumento na resistência da glote à corrente de ar (FANT, 1981; LADEFOGED; MCKINNEY, 1963; ROTHENBERG, 1973). Contudo, independentemente da ação da pressão subglótica, F0 também é influenciada pela tensão das cordas vocais. Dessa forma, a articulação na região laríngea poderia desempenhar algum papel na diferenciação de sílabas acentuadas e não acentuadas. Baseando-nos nessa possível relação acústico-fonatória, optamos por utilizar a eletroglotografia em nosso experimento, que será apresentado no cap. 6. Aumentos em F0 e, indiretamente, intensidade, poderiam ser identificados na forma de alterações do sinal glótico.

Nesta seção, procuramos apresentar resultados de estudos sobre a fala que mostram a importância de informações acústicas e fonatórias para a compreensão de fenômenos linguísticos e nos quais nos basearemos para analisar o acento. Apontamos evidências de que propriedades acústicas envolvidas na produção do acento poderiam ser relacionadas à articulação no nível glótico, o que nos motivou a utilizar em nossa coleta de dados um eletroglotógrafo, equipamento que permite a avaliação dos ciclos de vibração das pregas vocais.

Na seção seguinte passaremos a considerar estudos que tratam especificamente do acento no português, apontando suas contribuições para a descrição do fenômeno, bem como aspectos específicos decorrentes do modelo teórico adotado em cada estudo.

### 3.4 ESTUDOS SOBRE O ACENTO LEXICAL NO PORTUGUÊS

O acento lexical, no PB, se restringe a uma janela de três sílabas, ao fim da palavra, formando oxítonas, paroxítonas ou proparoxítonas (cf. BISOL, 1999; CAMARA JÚNIOR, 1970). Essa janela seria expandida para a quarta sílaba apenas em caso de epêntese (e.g. ['tekíniku], *técnico*, vogal epentética sublinhada), no sistema não-verbal, e de procliticização (e.g. *falávamos-te*), no sistema verbal.

Palavras paroxítonas prevalecem numericamente sobre proparoxítonas e oxítonas, ou seja, dentre os três padrões acentuais possíveis na língua, o paroxítono é o mais frequente (ALBANO, 2001). Este padrão seria responsável por imprimir o ritmo característico do português e do espanhol, em comparação com o acento predominantemente proparoxítono do italiano e exclusivamente oxítono do francês (CAMARA JÚNIOR, 1970, p. 55).

Como sugere Câmara Júnior (1970, p. 54), no PB, o acento tem função demarcativa, servindo para delimitar os vocábulos fonológicos, culminativa, por servir como âncora para a entoação, e distintiva, sendo responsável pelo contraste lexical em palavras como *sábia*, *sabia* e *sabiá*. Além disso, como se pode depreender dos exemplos anteriores, o acento tem também a função de contrastar classes lexicais (LEE, 1995, p. 141), o que recebe suporte do fato de não ser fácil encontrar pares mínimos de acento dentro da mesma classe de palavras – como *cara* x *cará*, no caso dos nomes, e *comeram* x *comerão*, no caso dos verbos –, fato que geralmente não é explorado pelas análises do acento no PB.

São vários os estudos que tratam da acentuação no PB (cf. Araújo (2007) para um apanhado de várias dessas análises). Apesar de estarem fundamentados em diferentes modelos teóricos, a descrição dos fatos linguísticos neles apresentada é razoavelmente equivalente. A maioria deles aponta para quatro aspectos fundamentais do acento:

- a) tem extensão abrangente, alcançando o domínio da palavra e grupos de palavras;
- b) envolve o estabelecimento de proeminência relativa contrastiva dentro do domínio de ocorrência;
- c) tem função de delimitar palavras, pois ocorre sempre em uma janela de três sílabas e interage com processos fonológicos como o de lenição silábica;
- d) interage com proeminências em outros níveis prosódicos, servindo como apoio para o ritmo e para a entoação.

Nos estudos sobre o acento lexical no português brasileiro, um dos pontos mais discutidos é em que medida ele é (1) especificado na estrutura lexical (“marcado no léxico”), (2) implementado por algoritmo de atribuição de acento ou (2) sensível ao peso silábico. Serão apresentados a seguir estudos que exemplificam cada posicionamento.

Camara Júnior (1970, p. 55), comprometido com a tradição estruturalista, defende que o acento no português é lexicalmente especificado, sendo que “sua posição não depende da estrutura fonêmica do vocábulo”.

Leite (1974) propõe uma análise gerativista do acento no PB. Avalia propostas anteriores, entre elas a de Camara Júnior (1970), e conclui que “stress and other phonetic properties are not independent phonetic characteristics of an utterance, dissociated from grammatical and syntactical considerations. They are the result of the application of a set of rules.” (LEITE, 1974, p. 26). Propõe que a informação sobre o acento não estaria presente no léxico, mas seria atribuída à cadeia sonora por meio de regra específica. Ainda, uma regra levaria em conta a ocorrência de determinados sufixos, todas elas sujeitas a condições específicas de aplicação. Em sua análise, oxítonas terminadas em vogal são marcadas no léxico para o acento. A autora conclui ser necessário postular regras diferentes para o sistema verbal e nominal, uma vez que a regra de atribuição de acento se aplicaria diferentemente em cada um dos sistemas.

Baseando-se na Fonologia Métrica, Bisol (1992, p. 69) apresenta a seguinte regra para atribuição do acento no PB que, ao contrário da regra gerativista de Leite (1974), faz referência a constituintes silábicos<sup>9</sup>:

Regra do Acento Primário

Domínio: a palavra

- i. Atribua um asterisco (\*) à sílaba pesada final, i.e, sílaba de rima ramificada.
- ii. Nos demais casos, forme um constituinte binário (não iterativamente) com proeminência à esquerda, de tipo (\* .), junto à borda direita da palavra.

Esta regra básica de atribuição de acento irá acentuar sílabas pesadas<sup>10</sup> finais (como em *amor*) e, na ausência destas, acentuar a primeira sílaba do pé<sup>11</sup>

<sup>9</sup> A sílaba apresenta uma estrutura hierárquica interna composta por *onset* (ou ataque), seguido de núcleo e coda, sendo estes três elementos chamados de constituintes silábicos. O *onset* é ocupado pelas consoantes iniciais e não exerceria influência na atribuição do acento. O núcleo é o centro da sílaba, componente obrigatório, equivalendo ao pico de sonoridade que a caracteriza, sendo ocupado por vogais, semi-vogais e, em algumas línguas, certas consoantes contínuas. A coda é ocupada por consoantes pós-nucleares. Núcleo e coda podem ser agrupados em um constituinte superior, a rima. (cf. HOGG; MCCULLY, 1987, p. 36-38).

<sup>10</sup> *Sílabas pesadas* são aquelas em que a rima é complexa (ou seja, sílabas em que o núcleo é ocupado por ditongo ou vogal longa e/ou a coda é preenchida), contrapondo-se às *sílabas leves*, que apresentam rima simples (cf. HOGG; MCCULLY, 1987, p. 36-38).

<sup>11</sup> O pé é um dos constituintes prosódicos na proposta de Nespor e Vogel (1986, p. 11), sendo superior à sílaba e inferior à palavra prosódica. Os pés podem ser classificados de acordo com sua estrutura interna. O pé trocaico, representado como (s w), é composto de sílaba forte (s) seguida de sílaba fraca (w); o pé iâmbico, ao contrário, é composto de sílaba fraca seguida de sílaba forte, sendo representado como (w s). Nespor e Vogel (1986) adotam o modelo arbóreo de representação, com a notação (s) para sílaba forte e (w) para sílaba fraca. No modelo da grade métrica, utilizado, por exemplo, em Halle e Vergnaud (1987), adota-se a notação (\*) e (.), em

localizado junto à borda direita da palavra (como em *travessura*, que, dividida em pés, (trave)(ssura), tem o primeiro pé mais à direita, (ssura), com a estrutura (\* .) ). Portanto, a análise de Bisol (1992) considera a estrutura silábica determinante para a localização do acento e defende que o pé básico no português seria troqueu – representado como (\* .). A autora tenta dar conta das exceções à regra básica recorrendo à extrametricidade<sup>12</sup> (para palavras com sílaba pesada final átona (e.g. *lápiz*) e para todas as proparoxítonas) e à presença de uma consoante final abstrata, ou “invisível” (para oxítonas terminadas em sílaba leve (como em *jacaré*)). Além disso, explica a diferença entre verbos e nomes em termos da natureza da regra de atribuição de acento: cíclica, nos não-verbos, não-cíclica, nos verbos.

Das abordagens do acento no PB que se valem da Teoria da Otimalidade, podemos mencionar Wetzels (2006). O autor sugere que o sistema acentual no PB é misto: nos verbos, o acento tem base morfológica e é sensível às categorias de tempo, enquanto, nos não-verbos, o acento tem base prosódica e é sensível ao peso silábico. O autor assume que o pé, no PB, tanto em verbos quanto em não-verbos, e mesmo nos casos excepcionais, tem cabeça à esquerda e é alinhado com a borda direita da palavra. Analisando as formas verbais, propõe três restrições para identificar a localização da cabeça do pé, de acordo com o tempo verbal (foot head present, foot head past, foot head future). Em seguida, passando aos não-verbos, apresenta seis restrições que definiriam as propriedades básicas do acento no PB, ranqueando-as da seguinte forma (WETZELS, 2006, p. 44):

match-prosody (si-pl), weight, align-ft-r, prwd-r, trochee >> binarity (σ)

match-prosody (si-pl), align-ft-r, prwd-r >> weight

A maioria dos estudos fonológicos sobre o acento no PB se vale de teorias não lineares, com predomínio da Fonologia Métrica e da Fonologia Lexical. Tais estudos, contudo, encontram problemas teóricos e empíricos, muitas vezes decorrentes de restrições dos modelos adotados – e.g. o frequente emprego de explicações *ad hoc* e artifícios teóricos, como extrametricidade e marcação no léxico/representação subjacente, para tratar casos que não se enquadram na análise. Ainda permanece em aberto o debate sobre em que medida o acento é especificado no léxico, implementado por algoritmo ou sensível ao peso silábico. Este estudo pretende contribuir com a compreensão dos mecanismos de atribuição do acento, adotando uma perspectiva teórica inovadora com relação às apresentadas anteriormente.

---

lugar de (s) e (w). No presente trabalho, não adotaremos a noção de pé, que é controversa.

<sup>12</sup> A extrametricidade é a mecanismo da Fonologia Métrica, proposto inicialmente por Liberman e Prince (1977), para lidar com dados que não se encaixavam perfeitamente nas regras. Funciona de modo a tornar as unidades classificadas como extramétricas invisíveis às regras de construção de constituintes (HALLE; VERGNAUD, 1987, p. 18).



### 3.5 CARACTERÍSTICAS ACÚSTICAS E FONATÓRIAS DO ACENTO NO PB

Fernandes (1977), avaliando as propriedades acústicas do acento no português na variedade paulista, identificou como relevantes a duração e em menor grau a frequência, seguida da intensidade. Ferreira Neto (2007), em estudo também sobre a variedade paulista, avaliou os seguintes parâmetros acústicos, procurando relacioná-los às funções distintiva e contrastiva do acento: intensidade, duração, frequência fundamental e qualidade vocálica (F1 e F2). O autor concluiu que a qualidade vocálica não bastaria como fator para a segmentação das palavras, como previsto pela função demarcativa do acento. Encontrou uma fraca correlação (1) entre frequência e intensidade e (2) entre acento, frequência e intensidade, separadamente e em conjunto. Por exclusão, concluiu que a duração seria o principal correlato acústico do acento. Ferreira Neto, contudo, não contemplou em sua análise as diferenças entre sílabas acentuadas e não acentuadas em cada um dos três padrões acentuais (oxítonas, paroxítonas, proparoxítonas), o que pode ter levado a uma redução no poder explicativo de seus resultados.

Major (1985) testou a produção do acento lexical por um falante baiano, um mineiro e um paranaense. Encontrou na duração o correlato mais confiável do acento. O autor descartou intensidade e frequência como correlatos do acento no português e, por essa razão, só apresentou os dados referentes à duração. Alega que, em seus dados, intensidade e frequência apresentaram inconsistências e grande variabilidade:

*In some utterances, mingograms revealed that the intensity of the pretonic was approximately equal to that of the posttonic, though the posttonic was shorter; or that the intensity and pitch of the pretonic were approximately equal to that of the tonic, though the pretonic was shorter; or that the intensity and pitch of the posttonic were approximately equal to that of the tonic, though the posttonic was shorter. Thus the relative pitch and intensity of the syllables may vary considerably, while the durational ratios remain fairly constant. This indicates that the primary correlate of stress in BP is length. (MAJOR, 1985, p. 261).*

No dialeto carioca, Moraes (1998) relata que a manifestação do acento é complexa, uma vez que a combinação entre os parâmetros duração, frequência e intensidade é influenciada pelo contexto de ocorrência.

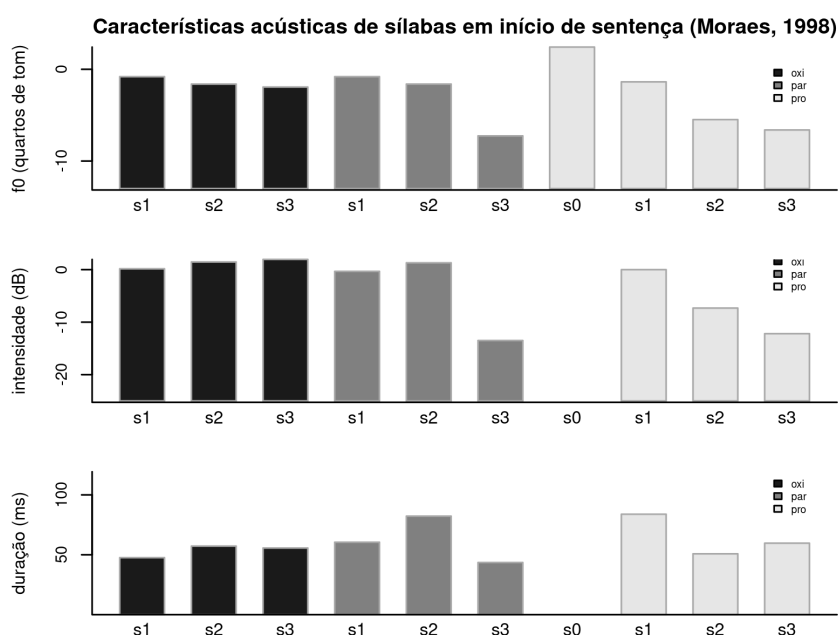
*In Brazilian Portuguese, acoustic correlates of lexical stress vary as a function of a word's position in a utterance and, less significantly, as a function of the word's stress pattern. When a word is found in a weak position (i.e., inside a prosodic group), stress prominence is expressed as a combination of intensity and duration. (MORAES, 1998, p. 180).*

Nas FIG. 6 e FIG. 7 são mostrados os resultados de Moraes (1998) para os correlatos acústicos do acento, respectivamente em início de sentença (posição fraca no domínio da entoação) e final de sentença (posição proeminente no domínio

da entoação). As classes s1, s2 e s3 representam respectivamente a antepenúltima, a penúltima e a última sílabas da palavra, enquanto a única s0 (linha de F0 e coluna de proparoxítonas) se refere à necessidade de especificar a sílaba anterior à antepenúltima: como mostram os dados de Moraes, o acento lexical corresponde a uma queda em F0 após a produção da sílaba acentuada. Para demonstrar tal padrão, faz-se necessário, portanto, especificar F0 na sílaba anterior à acentuada, o que, nas proparoxítonas, corresponde à sílaba anterior à antepenúltima – s0. A relação entre acento e intensidade, por outro lado, consiste antes na diferenciação das postônicas por meio da redução de intensidade. Em duração, o mesmo acontece, mas de forma menos pronunciada: as postônicas são mais curtas que as pretônicas, e ambas são mais curtas que as acentuadas.

**FIGURA 6** – Propriedades acústicas do acento lexical em início de sentença

a) F0 (quartos de tom); b) Intensidade (dB); c) Duração (cs)

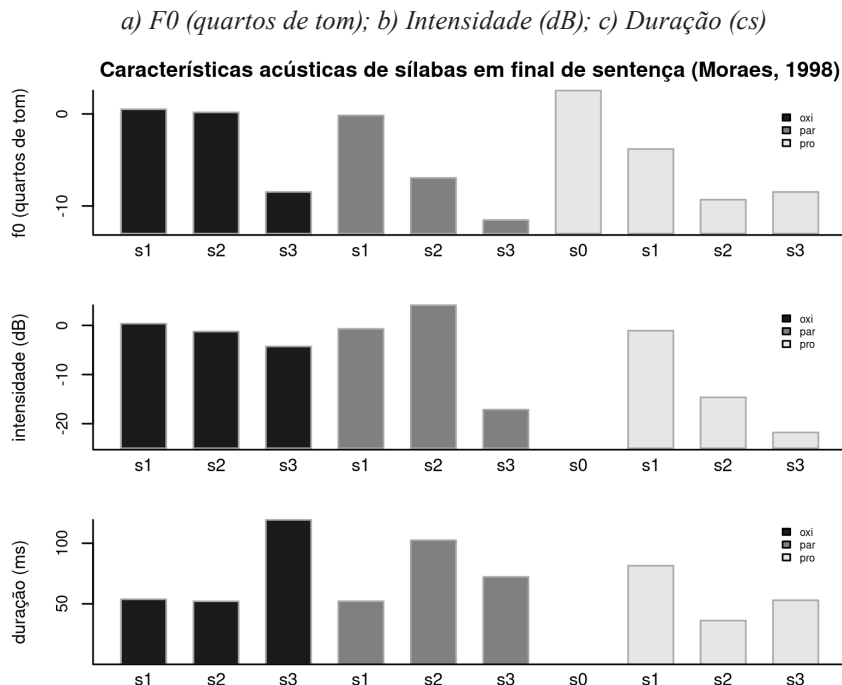


FONTE: MORAES, 1998, p. 180, adaptado. (Os dados de f0 que consistiam em uma curva ao longo da sílaba foram substituídos por um único valor: a média de f0 inicial e final naquela sílaba.)

Comparando as FIG. 6 e 7, o autor destaca ainda que, em início de sentença (FIG. 6), praticamente não há diferença acústica entre as três sílabas nas oxítonas, ao passo que, em posição final de sentença (FIG. 7), a sílaba acentuada passa a ser sinalizada pela queda em F0 e aumento significativo em duração. Moraes infere que a diferença entre as duas situações seria devida à sobreposição da

proeminência entoacional sobre a acentual, em fim de sentença. Comparando as FIG. 6 e 7, nota-se ainda um incremento na expressão acústica da proeminência: em posição final, os contrastes entre as sílabas acentuadas e não acentuadas são mais visíveis, principalmente em F0 e duração, intensidade em menor medida.

**FIGURA 7** – Propriedades acústicas do acento lexical em final de sentença



FONTE: MORAES, 1998, p. 181, adaptado. (Os dados de f0 que consistiam em uma curva ao longo da sílaba foram substituídos por um único valor: a média de f0 inicial e final naquela sílaba.)

Infelizmente, o português brasileiro não conta, ainda, com descrições acústicas tão variadas e abrangentes como outras línguas, em especial com relação ao acento. A maioria dos estudos sobre o acento no português assume uma perspectiva fonológica, não se ocupando das características acústicas do fenômeno. De fato, estudos acústicos em Prosódia, no PB, são majoritariamente devotados à entoação e ao ritmo.

Ademais, os estudos fonéticos sobre o fenômeno acentual, como os que apresentamos anteriormente, em geral consideram apenas propriedades acústicas e não avaliam os mecanismos fonatórios envolvidos na acentuação. Não foram encontrados, até o momento, estudos sistemáticos sobre a articulação do acento no português no nível glótico. Mesmo em outras línguas, tais estudos ainda são escassos, imagina-se, pela maior dificuldade atrelada à obtenção de dados diretos da função fonatória em comparação a dados acústicos.

Para alcançar o objetivo proposto de medição de propriedades acústicas e fonatórias envolvidas no acento, foi realizado, no presente estudo, um experimento de produção de fala. Avaliamos a variedade mineira. Tal experimento será relatado no capítulo 6.

A análise acústica realizada neste estudo contemplará as medidas de duração, intensidade e frequência fundamental, que a literatura aponta como relevantes para a realização do acento no português. Dessa forma, poderemos promover uma comparação com os dados de Fernandes (1977), Ferreira Neto (2007), Major (1985) e Moraes (1998), mencionados na seção 3.5.

Pretendemos ainda analisar propriedades articulatórias relacionadas à produção do acento no nível glótico, por meio de medidas referentes à dinâmica de fechamento das cordas vocais, obtidas por eletroglotografia. Pretendia-se adicionalmente avaliar a relação entre o acento do português e a movimentação corporal superior, por meio de medidas do deslocamento da mandíbula, da cabeça e da dinâmica dos gestos faciais, coletadas pelo Optotrak, mas a análise de tais dados excederia os limites de tempo deste estudo.

O QUADRO 2, a seguir, retoma os principais aspectos levantados pelos estudos fonológicos e acústicos do acento no PB, apresentados nas seções 3.4 e 3.5. Os aspectos foram divididos em três níveis de análise – funcional, físico e cognitivo –, com fins estritamente metodológicos, que nortearão o desenvolvimento da análise multirrepresentacional que será aqui proposta.

**QUADRO 2** – Aspectos a serem abordados na análise multirrepresentacional do acento proposta neste estudo

Nível de análise	Propriedades
Físico	Correlatos acústicos e articulatórios
Cognitivo	Previsibilidade por regra, por peso silábico ou especificação no léxico
Funcional	Interação entre morfologia, fonologia e prosódia
	Produtividade de padrões
	Contraste, demarcação e hierarquização de unidades prosódicas (função distintiva, demarcativa, culminativa)
	Diferenças referentes ao sistema nominal e verbal

Esses aspectos serão considerados sob uma perspectiva não baseada em traços e ou regras (ou qualquer mecanismo do tipo computacional, envolvendo unidades discretas e processamento serialista), mas antes em uma perspectiva multirrepresentacional, baseada no uso. O nível funcional será discutido no capítulo 5. Os níveis físico e cognitivo serão discutidos também no capítulo 5 e avaliados experimentalmente no capítulo 6.

